Available online at www.sciencedirect.com

SCIENCE DIRECT

Remote Sensing of Environment 98 (2005) 122 – 140

Remote Sensing of Environment

www.elsevier.com/locate/rse

An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation

P. Jeremy Werdell [a,*], Sean W. Bailey [b]

[a]Science Systems and Applications, Inc., NASA Goddard Space Flight Center, Greenbelt, Maryland, USA
[b]Futuretech Corporation, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Received 30 March 2005; received in revised form 27 May 2005; accepted 6 July 2005

Abstract

Global satellite ocean color instruments provide the scientific community a high-resolution means of studying the marine biosphere. Satellite data product validation and algorithm development activities both require the substantial accumulation of high-quality in-situ observations. The NASA Ocean Biology Processing Group maintains a local repository of in-situ marine bio-optical data, the SeaWiFS Bio-optical Archive and Storage System (SeaBASS), to facilitate their ocean color satellite validation analyses. Data were acquired from SeaBASS and used to compile a large set of coincident radiometric observations and phytoplankton pigment concentrations for use in bio-optical algorithm development. This new data set, the NASA bio-Optical Marine Algorithm Data set (NOMAD), includes over 3400 stations of spectral water-leaving radiances, surface irradiances, and diffuse downwelling attenuation coefficients, encompassing chlorophyll a concentrations ranging from 0.012 to 72.12 mg m^{-3}. Metadata, such as the date, time, and location of data collection, and ancillary data, including sea surface temperatures and water depths, accompany each record. This paper describes the assembly and evaluation of NOMAD, and further illustrates the broad geophysical range of stations incorporated into NOMAD.

Keywords: Ocean color; Satellite validation; Algorithm development; SeaBASS; SeaWiFS; MODIS; Bio-optics; Remote sensing; Water-leaving radiance; Chlorophyll

1. Introduction

The oceans contribute substantially to biological processes that help regulate the Earth's climate, yet, the geographic and temporal extent of such contributions is only partially understood. Coastal and inland waters support a diverse assortment of ecosystems, many of which possess commercial and ecological value. Still, a scientific understanding of the biological responses of these ecosystems to perturbations (e.g., climatic disturbances, erosion, and industrial pollution) has not been fully realized. The improved scientific understanding of such processes requires surmounting a considerable obstacle – the routine acquisition of high quality, globally distributed scientific observations – a feat near impossible using only ships and other marine platforms. Fortunately, satellite-borne ocean color instruments provide a regular, synoptic view of the productivity and variability of the Earth's oceans. A nadir-viewing, polar orbiting instrument with a 1-km^2 Earth footprint and 55° half-scan width, for example, observes an average of 15% of the ocean each day, and up to 50% over 4 days, after accounting for the effects of cloud coverage and contamination by excessive sun glint (Gregg et al., 1998). At such spatial and temporal scales, satellite ocean color instruments provide the scientific community a high-resolution means of studying the marine biosphere.

The color of seawater relies on the relative concentrations of optically active water-column constituents, including phytoplankton pigments, non-algal particulate and dissolved organic carbon, and water molecules themselves (Morel & Prieur, 1977). Chlorophyll *a*, the primary photosynthetic pigment in phytoplankton, absorbs relatively more blue and red light than green, and as its concentration increases, the spectrum of backscattered sunlight progressively shifts from blue to green (Yentsch, 1960). Satellite-borne ocean color instruments measure the spectral radiant flux emanating upward from the top of the Earth's atmosphere at discrete visible and near-infrared wavelengths. Algorithms are applied to these data to remove the contribution of the atmosphere from the signal (e.g., Gordon & Wang, 1994), thereby producing an estimate of the upwelling spectral radiant flux at the sea surface. The resulting water-leaving radiances, $L_w(\lambda)$, are in turn used to estimate a number of geophysical data parameters, such as the concentration of chlorophyll a, C_a, via the application of additional bio-optical algorithms (e.g., Carder et al., 1999; Maritorena et al., 2002; O'Reilly et al., 1998). The radiances and derived parameters, C_a in particular, are subsequently used to assess and monitor temporal changes in the marine ecosystem (e.g., Denman & Abbott, 1994; Siegel et al., 2002, and Subramaniam et al., 2002; Tomlinson et al., 2004) and to investigate the role of marine photosynthesis and net primary productivity in the

* Corresponding author.
E-mail addresses: jeremy.werdell@gsfc.nasa.gov (P.J. Werdell), sean.bailey@gsfc.nasa.gov (S.W. Bailey).

0034-4257/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2005.07.001

Earth's carbon budget (e.g., Antoine et al., 1996; Behrenfeld et al., 2001; Longhurst et al., 1995; Platt et al., 1991; Sarmiento et al., 2004).

Clark et al. (1970) successfully used data collected onboard high-altitude aircraft to relate the color of the ocean to its coincident chlorophyll a concentration. Less than a decade later, the Coastal Zone Color Scanner, launched onboard the National Aeronautics and Space Administration (NASA) Nimbus-7 spacecraft, provided the first (i.e., proof-of-concept) satellite ocean color data set (CZCS; Hovis et al., 1980). The extensive utility of this data for both coastal and open ocean research prompted the launch of a series of subsequent missions, including the Ocean Color and Temperature Sensor (OCTS; Tanii et al., 1991), the Sea-viewing Wide Field-of-view Sensor (SeaWiFS; McClain et al., 1998), the Moderate Resolution Imaging Spectroradiometer (MODIS; Esaias et al., 1998), and the Medium Resolution Imaging Spectrometer (MERIS; Rast & Bézy, 1995). All employ empirical algorithms to remotely estimate marine chlorophyll a concentrations.

The empirical relationship between satellite-derived ocean color and chlorophyll a concentrations has been studied for several decades (e.g., Clark, 1981; Gordon et al., 1980, 1983), culminating recently in the SeaWiFS Bio-optical Algorithm Mini-workshop (SeaBAM; Firestone & Hooker, 1998), an international collaboration whose primary goal was the identification of an operational chlorophyll a algorithm for SeaWiFS. A by-product of the SeaBAM effort was the compilation of a global, high-quality in-situ data set of coincident radiometric and chlorophyll a concentrations (O'Reilly et al., 1998). O'Reilly et al. (2000) expanded this SeaBAM data set (SBDS) prior to using it to define the current ocean chlorophyll a algorithms for OCTS, SeaWiFS, and MODIS. Today, the SBDS remains one of the largest in-situ data sets assembled and realizes continued value as a resource for the refinement and verification of bio-optical reflectance models (Carder et al., 1999; Maritorena et al., 2002; Tanaka et al., 2004; Yan et al., 2002). To our knowledge, it prevails as the most widely available public source of global, high-quality reflectance and chlorophyll a data for bio-optical algorithm development.

This notwithstanding, the observations in SBDS suffer from several deficiencies, such as a lack of associated metadata (e.g., date and time of collection and station latitude and longitude), making regional and temporal analyses impossible. Further, a mechanism for adding new observations or updating existing data has yet to be developed. Although more contemporary empirical algorithms have been successfully developed using regional data (Darecki & Stramski, 2004; D'Ortenzio et al., 2002; Garcia et al., 2005; Gohin et al., 2002; Kahru & Mitchell, 1999), they have rarely been verified on global scales, and generally consider only a local range of geophysical conditions. In principal, globally applied chlorophyll a algorithms produce the most widely acceptable results when developed using a cohesive global data set (e.g., Maritorena et al., 2002). The ongoing satellite missions (e.g., SeaWiFS, MODIS, and MERIS), and those scheduled to launch in the next decade (e.g., the Visible/Infrared Imager/Radiometer Suite (VIIRS) mission, part of the National Polar-orbiting Operational Enviromental Satellite System (NPOESS) Preparatory Project), require current in-situ data for their respective calibration and validation activities (Hooker & McClain, 2000; Werdell et al., 2003). Following, these missions, and the merger of their respective data products, also benefit from the use of bio-optical algorithms consistently developed with contemporary field data (Barnes et al., 2003). Further, as ocean color satellite data products (McClain et al., 2004) and processing utilities (Baith et al., 2001) have become widely available to an international community, the utility of and need for regionally tuned algorithms has increased significantly.

As such, we propose that a modern data set, with detailed metadata and an inherent mechanism for adding and updating data records, be considered prerequisite to adequately support present and future satellite ocean color missions. In this article, we describe the compilation and evaluation of such a data set, the NASA bio-Optical Marine Algorithm Data set (NOMAD), which is co-located with the NASA Ocean Biology Processing Group (OBPG) at Goddard Space Flight Center, in Greenbelt, Maryland, U.S.A. The data products included in NOMAD (Table 1) are used simultaneously for OBPG calibration and validation activities. Following the legacy of the SBDS, for remote sensing studies, these pertinent geophysical data products include spectral water-leaving radiances and surface irradiances (the ratio of which provides remote sensing reflectance), spectral column-averaged diffuse attenuation coefficients, and chlorophyll a concentrations. We also recorded weekly averaged sea surface temperature (optimum interpolation sea surface temperature; Reynolds et al., 2002) and station water depths (National Geophysical Data Center ETOPO2; Jakobsson et al., 2000; Smith & Sandwell, 1997) with each data record, but do not discuss either in this article.

5. Conclusions

NOMAD is a publicly available, global, high quality in-situ bio-optical data set for use in ocean color algorithm development and satellite data product validation activities. Current data products include coincident observations of water-leaving radiances and chlorophyll a concentrations, along with relevant metadata, such as date, time, and coordinates of data collection, and ancillary data, including sea surface temperatures and water depths. This existing suite of data products both encourages the improvement of existing global ocean color chlorophyll algorithms (O'Reilly et al., 2000), and facilitates the development of regional empirical algorithms. We plan on including inherent optical properties (e.g., spectral absorption and backscattering coefficients) in the near future to support the analysis and evaluation of semi-analytic algorithms (Carder et al., 1999; Garver & Siegel, 1997; Hoge & Lyon, 1996; Lee et al., 2002; Roesler & Perry, 1995). The radiometric and pigment profiles used in the development of NOMAD are publicly available in SeaBASS (Werdell & Bailey, 2002). The NOMAD data set is currently available online via two mechanisms, a digital text file, adhering to ASCII format, which includes the full merged bio-optical data set, and an Internet search engine that provides a means of limiting the data to specific data products, field campaigns, or date and location ranges. All are available via the SeaBASS Web site (http://seabass.gsfc.nasa.gov).